tients and 62 peritoneal-dialysis patients at Siriraj hospital. Utility scores and the correlation coefficient with KDQOL-36 were calculated. Percentages of respondents with the ceiling and floor effects were compared for each of the different measurement tools. **RESULTS:** Patient samples had a mean age of 60.20±14.84 years. Mean duration of dialysis were 7.44±5.42 years for hemodialysis patients, and 1.82±1.22 years for peritoneal-dialysis patients. The mean SF-6D score (0.783±0.164) was significantly higher than EQ-5D (UK: 0.752±0.309, Thai: 0.691±0.314), and VAS (0.666±0.196) scores. Most of the kidney specific dimensions were better correlated with SF-6D than EQ-5D (UK and Thai preference weight) and VAS scores. Ceiling effects were observed in the EQ-5D concerning both UK and Thai preference weight, due to the fact that the EQ-5D differentiates less in the better health states, whereas the floor effects were not clearly observed in any instrument tools. **CONCLUSIONS:** SF-6D presented better correlation with kidney specific scales while the responsiveness of EQ-5D utility scores was poor. One explanation might be a ''ceiling effect'' of the EQ-5D. These findings implied that SF-6D utility scores could reflect HRQoL status of dialysis patients better than EQ-5D and VAS.

PUK36

USING BOOTSTRAP CONFIDENCE INTERVALS TO COMPARE RELATIVE VALIDITY COEFFICIENTS: AN EXAMPLE WITH PRO MEASURES OF CHRONIC KIDNEY DISEASE (CKD) IMPACT

Deng N, Ware J
University of Massachusetts Medical School, Worcester, MA, USA

OBJECTIVES: To evaluate bootstrap techniques in comparing the validity of PRO measures in discriminating among CKD patients and responding to longitudinal changes. **METHODS:** The Kidney Disease Impact Scale (KDIS), CKD-specific legacy (KDQOL Burden, Symptom, and Effect) and generic health (SF-12) scales were administrated to 453 patients and re-administered to 110 patients after three months. ANOVA-based relative validity (RV) coefficients were used to compare how well each scale discriminated between three clinically-defined groups ordered in terms of severity (Dialysis > Stage 3-5 > Transplant), and how responsive each scale was to changes over time for self-evaluated Better, Same and Worse groups. Bootstrap was used to construct confidence intervals (CIs) to determine whether the differences in RVs were significant in comparisons between each scale and the best legacy measure - KDQOL Burden. Sample size, number of bootstrap iterations, and type of CIs were varied to evaluate their impacts on CI using real and artificial data. **RESULTS:** The sample size played a substantial role. 300 people for 3 groups were suggested as the minimum number to make meaningful comparisons between RVs using CI. Number of bootstrap replications (100 to 10,000) did not show an obvious effect on bootstrap standard error, although 300 showed improvement over 100 on CI. The bias-corrected and accelerated (BCa) type of CI was preferred for correcting both bias and skewness in bootstrap distribution and for producing narrower CIs. Using 95% CI and 300 sample size, differences in RVs were non-significant in comparisons with KDQOL-Burden (RV=1) for the following scales: SF-12 PCS (RV=.6), PF (RV=.7), RP (RV = .77), KDQOL-Effect (RV=.99), and KDIS (RV=1.13). **CONCLUSIONS:** Bootstrapping appears to be valuable in testing the significance of differences in the relative validity of these PRO measures from a statistical perspective. Samples of 100 per group compared and 300 bootstrap replications are recommended.

RESEARCH POSTER PRESENTATIONS – SESSION IV
RESEARCH ON METHODS STUDIES

RESEARCH ON METHODS – Clinical Outcomes Methods

PRM1

COMPLIANCE ON THE CONSOLIDATED STANDARDS OF REPORTING TRIALS (CONSORT) GUIDELINES IN RANDOMIZED CONTROLLED TRIALS

Godwin OP[1], Dyson B[1], Park SY[1], Lee E[2]
[1]*Howard University Center for Minority Health Services Research, College of Pharmacy, Washington, DC, USA,* [2]*Howard University, Washington, DC, USA*

OBJECTIVES: The Consolidated Standards of Reporting Trials (CONSORT) statement was published in 2001 and updated in 2010, strongly recommended the use of CONSORT diagram to report the flow of participants through each stages of the trial. This study was conducted to describe the level of compliance of the published clinical trial in following the CONSORT recommendations and to estimate prevalence of the compliance. **METHODS:** A systematic literature search of all randomized controlled trials of anti-infectious agents published in the top 10 general medicine journals and top 5 infectious disease journals published in 2010. The journals include: The New England Journal of Medicine, Journal of the American Medical Association, British Medical Journal (Clinical Research Ed), Archives of Internal Medicine, PloS Medicine, Annals of Internal Medicine, Clinical Infectious Diseases, the Journal of Infectious Diseases, the Lancet Infectious Diseases, AIDS, Emerging Infectious Diseases Journal, Annual Review of Medicine, Canadian Medical Association Journal, and Annals of Medicine Journal. Each article was reviewed by two independent investigators based on the reporting criteria recommended by the CONSORT statement. Exclusion criteria included non-randomized control studies, and studies not including intervention or control group. **RESULTS:** The study identified 129 published articles using explicit criteria on Medline search. A total of 73 randomized controlled trials met the inclusion criteria. Of 73 studies, 55 (75.34%) articles included the CONSORT diagram. A comprehensive depiction of the CONSORT guidelines will be made and detail descriptions on the compliances will be presented by journal types during the presentation. **CONCLUSIONS:** Randomized controlled trials published in the top 10 general medicine journals and the top 5 infectious diseases journals in 2010 contain significant deficiencies in reporting the CONSORT flow chart. The clarity and the completeness of a study could be improved if the CONSORT statement is followed as prescribed.

PRM2

NETWORK META-ANALYSIS OF INDIVIDUAL AND AGGREGATE LEVEL DATA

Jansen JP[1], Cope S[2]
[1]*Mapi Consultancy / Tufts University School of Medicine, Boston, MA, USA,* [2]*Mapi Consultancy, Boston, MA, USA*

OBJECTIVES: Network meta-analysis is often performed with aggregate level data (AD). A challenge with meta-regression models using AD is that the association between a patient level covariate and relative treatment effects of the compared interventions at the study level may not reflect the individual level effect-modification. In this paper, non-linear network meta-analysis models for combining individual patient data (IPD) and AD are presented to reduce bias and uncertainty of treatment effects in the presence of heterogeneity due to patient characteristics. **METHODS:** The first method uses the same model form for IPD and AD. With the second method, the model for AD is obtained by integrating an underlying IPD model over the joint within-study distribution of covariates. With a simple simulation study the two modeling approaches are compared. **RESULTS:** Having IPD for a subset of studies improves estimation of treatment effects with network meta-analysis in the presence of patient level heterogeneity and inconsistency. Of the two proposed non-linear models for combining IPD and AD, the second approach seems less affected by bias. Additional studies, however, are needed to assess the value of both methods. **CONCLUSIONS:** Overall, for network meta-analysis it is recommended to use IPD when available, rather than treating all studies as AD.

PRM3

THE ENSEMBLE MINIMUM DATASET: A NEW INSTRUMENT TO EXPLORE HETEROGENEITY OF TREATMENT EFFECT

Brenneman SK[1], Brekke L[1], Bancroft TJ[1], Shen W[2], Paczkowski R[2], Berger M[1], Kaplan SH[3], Greenfield S[3], Buesching DP[2]
[1]*OptumInsight, Eden Prairie, MN, USA,* [2]*Eli Lilly and Company, Inc., Indianapolis, IN, USA,* [3]*University of California Irvine, Irvine, CA, USA*

OBJECTIVES: To develop an instrument that identifies patient groups likely to have differing responses to treatment, we tested candidate measures thought to discriminate differences among patients in 4 disease cohorts: type 2 diabetes (T2D), knee osteoarthritis (OA), ischemic heart disease (IHD) and heart failure (HF). **METHODS:** Eligible patients identified from claims data were sent a survey including 17 scales hypothesized to comprise 4 domains (health profile, personality, behavior, life context). Proxies for treatment response were patient-reported global impression of disease severity (PGIS), global impression of improvement (PGII), and administrative claims health care utilization (HCU). Variability (SD) and internal consistency (Cronbach's alpha) of the scales were examined, as was discriminant validity against strata of PGIS, PGII and HCU. Conceptual overlap, correlations among scales, and factor loading within and across domains were examined. Scales with desirable properties were included in the final instrument. Discriminant validity of proposed domains was analyzed by ANOVA adjusted for age and gender. Multiple regression models were used to assess the associations between the proposed domains and outcomes. **RESULTS:** A total of 723 T2D patients, 682 knee OA patients, 632 IHD patients, and 588 HF patients completed the survey. The initial instrument was refined to 7 scales across 3 domains. The health profile domain significantly discriminated 100% of the strata across disease cohorts (each $P<0.001$). Personality and behavior domains also discriminated strata well (75% and 50%, respectively). Alone, the health profile significantly discriminated strata across disease cohorts in multivariate analyses (each $P<0.001$). In models including all 3 domains, the health profile remained the most important contributor. **CONCLUSIONS:** The final ENSEMBLE MDS instrument discriminated among patients with varied diseases; the health profile provided much of the ability to discriminate. Further research is needed to assess the instrument's potential to predict health state changes due to trial interventions.

PRM4

ENHANCING THE HEALTH ECONOMIC VALUE OF RETROSPECTIVE AND PROSPECTIVE REAL-WORLD STUDIES WITH PHARMACOGENOMIC TESTING: OPPORTUNITIES AND CHALLENGES ASSOCIATED WITH AN INTEGRATED PERSONALIZED MEDICINE APPROACH

Payne KA[1], Frueh FW[2], Sohal J[3]
[1]*United BioSource Corporation, Dorval, QC, Canada,* [2]*Medco Health Solutions, Inc., Franklin Lakes , NJ, USA,* [3]*United BioSource Corporation, Hammersmith, UK*

OBJECTIVES: A better understanding of a patient's genetic make-up through pharmacogenomic testing can help achieve improved and more predictable patient outcomes, often at equal or lower total treatment cost. Stakeholders including physicians, payers and patients alike can benefit from real-world data that identify, a priori, the sub-groups of patients for whom treatments are likely to be more cost-effective. **METHODS:** Retrospective and prospective case study designs within which pharmacogenomic testing has been integrated are presented. Design parameters are described and opportunities and challenges alongside strategies for resolution are delineated. **RESULTS:** As the genetic make-up of a patient does not change, pharmacogenomic testing can be done at any point in time and paired with historical and/or newly collected patient level data. Retrospective studies are highly efficient as they do not require costly longitudinal follow-up, whereas prospective studies including registries offer the opportunity to augment pharmacogenomic and other study data with patient and physician reported outcomes not otherwise available in the medical chart. Main challenges associated with either approach include optimizing the patient informed consent process, streamlining the logistics associated with pharmacogenomic testing and storage in the usual

care environment, and data analytics. **CONCLUSIONS:** The integration of pharmacogenomic testing with real-world studies offers an important opportunity to identify sub-groups of patients for whom treatment is more effective in terms of clinical, and safety outcomes. Alongside resource utilization and cost of care data, this evidence can be used to populate cost-effectiveness and other health economic analyses to inform physician and payer decision-making.

PRM5

VALIDITY OF REQUIRING A MINIMUM DURATION OF POST-INDEX ENROLLMENT IN RETROSPECTIVE DATABASE STUDIES

Lanes SF
United BioSource Corporation, Lexington , MA, USA

OBJECTIVES: Retrospective database studies commonly use an inclusion criterion requiring that subjects have a minimum duration of post-index enrollment (i.e., follow-up). Such a criterion can simplify analysis and facilitate computation of annual costs. In clinical trials, however, similar strategies, such as analyses restricted to subjects who completed follow-up ("complete case analysis"), are seen as problematic because reasons for discontinuation may be related to study endpoints (i.e., informative censoring). **METHODS:** We reviewed methodologic literature and we used a health insurance claims database to evaluate the impact on health care utilization and costs of excluding subjects lost to follow-up. **RESULTS:** Excluding from analysis subjects with incomplete follow-up may be valid if patients are missing at random. Unfortunately, this assumption can rarely be verified because endpoints are usually unknown for patients who are lost to follow-up. In an insurance claims database, an inclusion criterion requiring one year of follow-up decreased health care utilization and average annual costs by 8% for a random sample of subjects, and by 17% among subjects with a serious illness. **CONCLUSIONS:** Subjects are lost to follow-up in both clinical trials and retrospective database studies (e.g., by exiting the database). Study populations should not be defined in such a way as to exclude subjects lost to follow-up; instead, subjects lost to follow-up should be considered as a missing data problem. In retrospective database studies, just as in clinical trials, if endpoints among subjects lost to follow-up differ from endpoints among subjects remaining in the database, restricting analysis to patients with minimum durations of follow-up can distort outcomes and economic evaluations. Subjects lost to follow-up in automated databases should be described and evaluated for evidence of informative censoring, and analyzed using strategies appropriate for missing data, such as multiple imputation methods.

PRM6

ARE YOU COUNTING PRESCRIPTION MEDICATIONS UTILIZATION CORRECTLY?

Athavale AS[1], Banahan BFI[1], Hardwick SP[2], Clark JP[2]
[1]*University of Mississippi, University, MS, USA,* [2]*Mississippi Division of Medicaid, Jackson, MS, USA*

OBJECTIVES: To evaluate the potential for duplicate counting of prescription medication utilization for products that are billed through medical and prescription claims. **METHODS:** A retrospective cross-sectional descriptive study was conducted using the 2008 – 2011 Mississippi Medicaid data. Medical claims (MCs) with J-codes for injectable medications were identified from MC files. Prescription claims (PCs) for the corresponding beneficiaries were extracted from PC data for all NDCs associated with the J-codes identified. These two sets of claims were stacked to obtain a denominator file. Potential duplicate counts were identified by pairing MCs and PCs for the same beneficiary and drug where the PC service date was within 7 days of the MC service date. The Medicare maximum allowable cost was identified for the J-code in each potential duplicate count situation. Criteria of the MC being 80+% of the maximum allowable cost for one J-code unit and the MC paid amount being 80+% of the corresponding PC paid amount were used to evaluate which pairs might be actual duplicate counts. **RESULTS:** Out of 1,813,251 claims identified in the denominator file, 1443 drug events were considered to be potential duplicate counts (0.08%). These claims were associated with 849 Medicaid enrollees. For 89% of the pairs, the MC paid amount was 80+% of the allowable J-code unit cost and 37% were 80+% of the corresponding PC paid amount. Using a combination of these criteria, it was estimated that at least 47% of the pairs were likely to be duplicate counts and that a large portion of the other pairs might be duplicate counts. **CONCLUSIONS:** Researchers need to use caution when counting medication events for products reimbursed as MCs and PCs. The error from over-counting at the population level should be small, but could have significant impact on utilization and adherence estimates for individual patients.

PRM7

A REVIEW AND APPLIED COMPARISON OF META-ANALYSIS TECHNIQUES

Lewis-Beck C[1], Baser E[2], Baser O[3]
[1]*STATinMED Research, Ann Arbor, MI, USA,* [2]*STATinMED Research, Istanbul, Turkey,* [3]*STATinMED Research/The University of Michigan, Ann Arbor, MI, USA*

BACKGROUND: Numerous assumptions and techniques are associated with performing meta-analysis. While some overall structural guidelines and recommended practices exist, there are very few papers that compare meta-analysis techniques in application. **OBJECTIVES:** To review primary meta-analysis methods and their assumptions, and apply various meta techniques to data and compare the results. **METHODS:** There are currently a myriad of meta-analysis techniques available. We started the study with a review of fixed effects models, which is the most basic technique that assumes homogeneity in treatment effect across studies. We then explored random effect models and meta regression. Each of these techniques models treatment heterogeneity. Other more advanced techniques examined included mixed treatment comparisons (MTC) and Bayesian approaches. **RESULTS:** Estimates of treatment effect differed depending on the meta technique applied. When a fixed effect model was applied to estimate the effect of a vaccination against tuberculosis, the log odds ratio was -0.436 (confidence interval [CI: -0.528, -0.344]). After testing for heterogeneity and fitting a random effects model, the estimate was reduced to -0.741 (CI [-1.12, -0.352]), and the CI became wider. When covariates were added to the model to explain the heterogeneity, the treatment effect was reduced even further. Additional techniques were applied as well, such as Bayesian MTC. **CONCLUSIONS:** Results from meta-analysis are sensitive to the studies selected, in addition to the methodology applied. To ensure that proper techniques are used, it is critical to estimate an unbiased outcome.

RESEARCH ON METHODS – COST METHODS

PRM8

GENERAL TRANSFERABILITY OF MODEL-BASED ECONOMIC EVALUATIONS

Carswell C, McWilliams P, Williamson KA, Faulds D
Adis, Auckland, New Zealand

OBJECTIVES: Economic evaluations of drug therapy are important, but time consuming and costly. Analyses that are easily transferable (i.e. adjustable to a different jurisdiction without completely rebuilding the model) may potentially save time and resources. We aimed to develop a tool to assess and summarize the general transferability of model-based analyses. **METHODS:** Medline was searched for literature on transferability published between 2002 and June 2011. Existing checklists for economic evaluations were adapted to create a checklist of 16 key factors to assess the general transferability of model-based analyses. This tool was used to score 11 recently published economic evaluations and identify how well specific factors were addressed. **RESULTS:** Transferability scores of the selected papers ranged from 53–91%, illustrating the wide variability in the quality of reporting. Across all studies, the least well addressed transferability factors included the discussion of the generalizability of the study results (lacking or incomplete in all studies), adequate description of resources and costs employed in the analysis (particularly separate reporting of resource use and unit costs), and adequate descriptions of the method and/or populations used to derive utility values. The best addressed transferability factors included those relating to country, currency and discount rates. Even if studies scored highly overall, it may still be difficult to transfer the findings to a different setting if they failed to report insufficient detail on one or two key parameters. **CONCLUSIONS:** The general transferability of a model-based economic evaluation from one country or jurisdiction to another can be quickly assessed by the application of a simple checklist of key transferability factors. It is important that authors ensure that they report their economic analysis in a detailed and transparent fashion.

PRM9

RELEVANCE AND QUALITY OF THE PHARMACOECONOMIC LITERATURE OF FDA RECENTLY APPROVED DRUGS: A SYSTEMATIC REVIEW

Woersching AL, Raisch DW, Borrego M
University of New Mexico College of Pharmacy, Albuquerque, NM, USA

OBJECTIVES: To perform a systematic literature review of pharmacoeconomic (PE) publications considering recent United States (US) Food and Drug Administration (FDA) new molecular entity and new biologic license approvals (NMEs/NBLs). The review investigated publication quality and US relevance. **METHODS:** MEDLINE and the United Kingdom National Health Service Economic Evaluation Database were searched. Included publications considered 2008-2009 NMEs/NBLs in original PE evaluations. In addition to general characteristics, each publication was evaluated using the Quality of Health Economic Studies (QHES) Instrument. The correlation between QHES scores and the 2010 Thomson Reuters five-year journal Impact Factor (IF-5y) was calculated. Median QHES score differences were compared (Mann-Whitney U) by study characteristics (yes/no): US context, academic first author, pharmaceutical manufacturer funding (PMF), and declared author independence. **RESULTS:** From 115 unique search results, 31 met inclusion criteria. Of fifty 2008-2009 NMEs/NBLs, 36% had PE publications, with 81% considering the approval indication and 61% published post-approval. A US context was assessed in 35% of publications. PMF was present in 68% of publications, comprising manufacturers marketing either the NME/NBL, 90%, or a comparator, 10%. Time (mean±standard deviation (S.D.)) since FDA approval was 21.9±8.8 months until ePublication and 15.3±9.0 months until journal submission. Median and mean±S.D. QHES score were 78 and 73.3±16.4, respectively. Publications most often satisfied QHES items regarding uncertainty (5) and incremental analysis (6) (94% each). Justifiying the chosen model (13) and discussing biases (14) were satisfied least often (38% each). The IF-5y (mean= 3.46, S.D.= 3.37) was not correlated with QHES score (Pearson r=0.095, p=0.636). QHES scores were not-significantly different ($p>0.05$) for any study characteristics. **CONCLUSIONS:** QHES scores indicate PE studies of recent NMEs/NBLs are high quality, although US relevance is imperfect: few publications assessed a US context; some did not consider the approval indication; publication lags delay PE evidence availability; and most publications have PMF.

PRM10

THE BEHAVIORAL ECONOMICS OF THE MINIMALLY IMPORTANT DIFFERENCE

Suh JK, Doctor J
USC School of Pharmacy, Los Angeles, CA, USA

OBJECTIVES: To study whether the minimally important differences (MIDs) values outcomes based on the behavioral economic theory. **METHODS:** We studied the behavior of individuals discriminating minimally important differences (MIDs), a method that identifies the change in a health measure necessary for a patient to discriminate an improvement. The behavioral theory predicts that discrimination of a quantity is governed by Weber's Law: If a quantity is increased by some factor,